J. Math. Anal. Appl. 356 (2009) 355–362

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Cubically convergent methods for selecting the regularization parameters in linear inverse problems ☆

Yongkui Zou, Linjun Wang, Ran Zhang *

School of Mathematics, Jilin University, Changchun 130012, PR China

A R T I C L E I N F O

Article history:
Received 26 October 2008
Available online 11 March 2009
Submitted by P. Sacks

Keywords:
Inverse problem
Morozov's discrepancy principle
Regularization parameters
Iterative method
Cubic convergence

A B S T R A C T

We present three cubically convergent methods for choosing the regularization parameters in linear inverse problems. The detailed algorithms are given and the convergence rates are estimated. Our basic tools are Tikhonov regularization and Morozov's discrepancy principle. We prove that, in comparison with the standard Newton method, the computational costs for our cubically convergent methods are nearly the same, but the number of iteration steps is even less. Numerical experiments for an elliptic boundary value problem illustrate the efficiency of the proposed algorithms.

1. Introduction

Inverse problems arise in a wide spectrum of applications in fields ranging from engineering to scientific computation [4,7,8,17,18]. In these problems one often has to solve operator equations of the first kind, which are usually ill-posed in the sense of Hadamard [4,18]. It means that the hardest issue in the numerical computation of inverse problems is the instability of the solution with respect to the noise from the observation data; that is, small perturbations of the observation data may lead to large changes on the considered solution. Thus to ensure a feasible and stable numerical approximation solution, it is necessary to employ some kind of regularization method, and hence it is necessary to develop appropriate strategies for choosing the regularization parameters. In practice, the effectiveness of a regularization method depends strongly on the choice of a good regularization parameter.

So far, a significant amount of research work has focused on the development of appropriate strategies for selecting the regularization parameters (see [1,2,4,10,16] and references therein), while less work has been carried out on the numerical realization of such strategies. Kunisch and Zou [14] pointed out that very few of these strategies were utilized for practical applications.

Motivated by the technique developed by Hebden in [3] for quasi-Newton methods, Ito and Kunisch [11] proposed a model function approach with four parameters to solve some nonlinear parameter identification problems. Kunisch and Zou [14] introduced two second-order iterative methods and a two-parameter model function method to choose reasonable regularization parameters in the Tikhonov regularization formulation of linear inverse problems. The basic tool is based upon the well-known Morozov's discrepancy principle [4,15,16] and the damped Morozov's discrepancy principle [5,6,11,13–15].

☆ This work is partly supported by NSFC (Nos. 10671082, 10801062 and 10701039), the (postgraduate) program of 985 of Jilin University (No. 20080239), the Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education.

* Corresponding author. Fax: +86 431 85168310.
E-mail address: zhangran@jlu.edu.cn (R. Zhang).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2009.03.024

Xie and Zou [20] improved the method studied in [14], proposed a new model function iterative method, and updated the model parameters in a computationally more stable manner.

Also, in testing the Tikhonov regularization method for solving an inverse problem, one needs to test a large number of regularization parameters, and this can often be very time-consuming. A significant reduction in time and computational cost can be achieved if there is an iterative method which can yield a reasonable regularization parameter at acceptable computational costs. In general, the commonly used iterative methods for solving the discrepancy principle are Newton method with quadratic convergence and quasi-Newton method with superlinear convergence (see [14]). Based on this, Wang and Xiao [19] promoted a cubically convergent algorithm based on Tikhonov regularization by using Taylor expansion.

The aim of this paper is to present some further work on the parameter selection strategies. We derive a general third-order iterative method for choosing the regularization parameters in linear inverse problems. In fact, the third-order method could be well used for most of the a posteriori parameter selection strategies.

In the following, we first review some basic notations and useful results [14]. Based upon the Tikhonov regularization method and Morozov's discrepancy principles we deduce a nonlinear equation that the regularization parameter has to satisfy [14]. We then apply three different iterative methods to these nonlinear equations and prove that these iterative methods all exhibit third-order convergence. We then compare the computational costs for each iteration step with the one of the standard Newton method. If n denotes the dimension of the corresponding finite element space, we find that our cubically convergent methods need $\frac{1}{6}n^3 + \frac{5}{2}n^2 + \frac{41}{6}n$ multiplications and divisions for each iteration step, which is just $\frac{1}{2}n^2 + \frac{21}{6}n$ more than the number for the standard Newton method. Therefore, when n becomes large, the computational costs of our cubically convergent methods are nearly the same with the one of Newton method. In addition, numerical experiments for elliptic inverse problems also illustrate the efficiency of the proposed algorithms. The numerical computations suggest that for large n our cubically convergent methods are faster than Newton method for computing the regularization parameters.

Consider a linear ill-posed inverse problem of the following form:

$$Kx = y, \tag{1.1}$$

where $K : X \to Y$ is a bounded linear operator from the parameter space X to the observation space Y. $\|\cdot\|_X$ and $\|\cdot\|_Y$ denote the norms in the Hilbert spaces X and Y respectively ($(\cdot,\cdot)_X$ and $(\cdot,\cdot)_Y$ denote the corresponding inner products). Here we call the problem (1.1) ill-posed in the sense that the solution of problem (1.1) does not depend continuously on the right-hand side observation data which are often corrupted by errors.

Assume that y^δ denotes the noisy observation data, and

$$\|y - y^\delta\|_Y \leqslant \delta,$$

where $\delta > 0$ is the amplitude of the given noise.

In order to approximate the ill-posed problem (1.1) by a well-posed one and compute approximation solution in a stable way, regularization methods are widely applied. Among these methods, the Tikhonov approach is the best known. The idea underlying the Tikhonov approach is to replace the least squares problem $\min_{x\in X} \|Kx - y^\delta\|_Y$ by a modified optimal problem of the form

$$\min_{x\in X} J(x,\alpha) = \frac{1}{2}\|Kx - y^\delta\|_Y^2 + \frac{\alpha}{2}\|x\|_X^2, \tag{1.2}$$

where $\alpha > 0$ is the regularization parameter.

We end this section with a well known existence result for problem (1.2). Let $K^* : Y \to X$ denote the adjoint operator of K. For any given α, let $x(\alpha)$ be the solution of problem (1.2).

Lemma 1.1. *(See [12].) For any $\alpha > 0$, there exists a unique solution $x(\alpha)$ to the minimization problem* (1.2). *It is characterized as the solution of the system*

$$K^*Kx + \alpha x = K^*y^\delta, \tag{1.3}$$

or, in variational form,

$$(Kx, Kg)_Y + \alpha(x, g)_X = \left(y^\delta, Kg\right)_Y, \quad \textit{for all } g \in X. \tag{1.4}$$

2. Iterative realization of parameter selection strategies

In this section, we state some useful lemmas and describe the numerical realization of the parameter selection strategies. First we introduce a lemma on the differentiability of the solution $x(\alpha)$ with respect to the parameter α.

Lemma 2.1. *(See [14].) The function $x(\alpha)$ is infinitely differentiable at every $\alpha > 0$ and its nth order derivative $x^{(n)}(\alpha) \in X (n \geqslant 1)$ is the unique solution of the following variational problem: find $\omega \in X$ such that*

$$(K\omega, Kg)_Y + \alpha(\omega, g)_X = -n\left(x^{(n-1)}(\alpha), g\right)_X, \quad \textit{for all } g \in X. \tag{2.1}$$

Define a function $F(\alpha)$ by

$$F(\alpha)=\frac{1}{2}\left\|Kx(\alpha)-y^{\delta}\right\|_Y^2+\frac{\alpha}{2}\left\|x(\alpha)\right\|_X^2.$$

We have the following lemma.

Lemma 2.2. *For any $\alpha>0$, the first, second and third derivative of the function $F(\alpha)$ with respect to α are respectively given by*

$$F'(\alpha)=\frac{1}{2}\left\|x(\alpha)\right\|_X^2, \tag{2.2}$$

$$F''(\alpha)=\big(x(\alpha),x'(\alpha)\big)_X, \tag{2.3}$$

$$F'''(\alpha)=\big(x'(\alpha),x'(\alpha)\big)_X+\big(x(\alpha),x''(\alpha)\big)_X. \tag{2.4}$$

Proof. The proof of formula (2.2), (2.3) is similar to the proof of Lemma 2.2 in [14]. We obtain Eq. (2.4) by directly differentiating the function $F''(\alpha)$ with respect to α. $\square$

Lemma 2.3. *(See [14].) Assume that $y^{\delta}\notin\ker K^*$. Then the non-negative function $F(\alpha)$ is strictly monotonically increasing and strictly concave.*

Next, we discuss the iterative realization of the parameter selection strategy. We will repeatedly use the expressions for the first, second and third derivative of the function $F(\alpha)$ given in Lemma 2.2, together with the fact that these derivatives can be obtained by solving the linear equation (2.1) in a stable manner if α is large. Our discussion is based on Morozov's discrepancy principle.

We first state a useful identity. Let $x(\alpha)$ be the unique minimizer to problem (1.2) for $\alpha>0$. According to Lemma 1.1, we have

$$K^*Kx(\alpha)+\alpha x(\alpha)=K^*y^{\delta}. \tag{2.5}$$

Differentiating both sides of Eq. (2.5) with respect to α, we obtain

$$K^*Kx'(\alpha)+\alpha x'(\alpha)+x(\alpha)=0. \tag{2.6}$$

Taking the inner product with $x(\alpha)$ in Eq. (2.6), we are led to the equation

$$\big(Kx'(\alpha),Kx(\alpha)\big)_Y+\big(x(\alpha),x(\alpha)\big)_X+\alpha\big(x'(\alpha),x(\alpha)\big)_X=0, \tag{2.7}$$

according to Lemma 2.2, we have

$$2F'(\alpha)+\alpha F''(\alpha)+\frac{1}{2}\frac{d}{d\alpha}\big(Kx(\alpha),Kx(\alpha)\big)_Y=0;$$

that is, for all $\alpha>0$,

$$\frac{d}{d\alpha}\left\{\alpha F'(\alpha)+F(\alpha)+\frac{1}{2}\big(Kx(\alpha),Kx(\alpha)\big)_Y\right\}=0.$$

Therefore,

$$2\alpha F'(\alpha)+2F(\alpha)+\big(Kx(\alpha),Kx(\alpha)\big)_Y=2C_0,$$

where C_0 is an integration constant.

The well-known Morozov discrepancy principle has been used for linear ill-posed problems to choose the regularization parameter. The principle is devoted to the posteriori choice of the regularization parameter α. The idea of the strategy is to choose the regularization parameter so that the residual $\|Kx(\alpha)-y^{\delta}\|_Y$ has the same error level as the observation data. That is, we require that α is determined by the equation

$$\left\|Kx(\alpha)-y^{\delta}\right\|_Y=\delta, \tag{2.8}$$

where δ is the observation error defined by

$$\delta=\left\|y-y^{\delta}\right\|_Y.$$

Throughout this paper we assume that $y^{\delta}\notin\ker K^*$. Then we can rewrite Eq. (2.8) as

$$F(\alpha)-\alpha F'(\alpha)=\frac{1}{2}\delta^2. \tag{2.9}$$

The next lemma presents the existence and uniqueness of the solutions of the exact Morozov equation (2.9).

Lemma 2.4. *(See [14].) If $F(0) < \frac{1}{2}\delta^2 \leqslant F(1)$, then there exists a unique solution $\alpha^* \in (0, 1]$ of the Morozov equation* (2.9).

In general, people use Newton method or the quasi-Newton method to solve the Morozov equation

$$G(\alpha) = F(\alpha) - \alpha F'(\alpha) - \frac{1}{2}\delta^2 = 0. \tag{2.10}$$

According to Lemma 2.2, we compute

$$G'(\alpha) = F'(\alpha) - F'(\alpha) - \alpha F''(\alpha) = -\alpha F''(\alpha) = -\alpha\big(x(\alpha), x'(\alpha)\big)_X.$$

Newton method for solving Eq. (2.10) can be described as follows: for a given initial guess α_0, generate the Newton sequence $\{\alpha_n\}_{n=0}^{\infty}$, by

$$\alpha_{n+1} = \alpha_n - \frac{G(\alpha_n)}{G'(\alpha_n)} = \alpha_n + \frac{G(\alpha_n)}{\alpha_n(x(\alpha_n), x'(\alpha_n))_X}.$$

The process of the computation involves the evaluation of $x'(\alpha)$ which is obtained by solving the following equation for v:

$$K^*Kv + \alpha v = -x(\alpha). \tag{2.11}$$

Obviously, the Newton method converges quadratically. However, we must solve both for $x(\alpha)$ and $x'(\alpha)$ at each iteration step, and this represents the major cost in the computational process.

In order to avoid solving Eq. (2.11), Kunisch and Zou [14] replaced $x'(\alpha_n)$ in the Newton method by the finite difference quotient

$$x_n(\alpha_n, \alpha_{n-1}) = \frac{x(\alpha_n) - x(\alpha_{n-1})}{\alpha_n - \alpha_{n-1}},$$

which leads to the quasi-Newton method.

Newton method and the quasi-Newton method for solving the discrepancy principle converge quadratically and superlinearly, respectively (see [14]). The implementation of these two methods involves a huge amount of computation in each iterative process for choosing a reasonable regularization parameter. In this paper, we apply a cubically convergent algorithm which needs nearly the same computational cost in each iteration step, but reduces the number of iteration steps.

We apply the following family of third-order methods [9] to compute the corresponding regularization parameter:

$$\alpha_{n+1} = \alpha_n - \left(1 + \frac{1}{2}L(\alpha_n)\big[1 - \beta L(\alpha_n)\big]^{-1}\right)\big[G'(\alpha_n)\big]^{-1} G(\alpha_n),$$

where $\beta \in [0, 1]$ and

$$L(\alpha_n) = G(\alpha_n)\big[G'(\alpha_n)\big]^{-2} G''(\alpha_n).$$

This family of methods includes the classical Chebyshev ($\beta = 0$), Halley ($\beta = \frac{1}{2}$), and Super-Halley ($\beta = 1$) methods.

In order to apply these methods, we need to evaluate an extra second-order derivative of $G''(\alpha_n)$. It follows from Lemma 2.2 that

$$G''(\alpha) = -\alpha\big[\big(x'(\alpha), x'(\alpha)\big)_X + \big(x(\alpha), x''(\alpha)\big)_X\big] - \big(x(\alpha), x'(\alpha)\big)_X.$$

By Lemmas 1.1 and 2.1, we can obtain $x(\alpha)$, $x'(\alpha)$ and $x''(\alpha)$ by solving the following equations, respectively:

$$(K^*K + \alpha I)x(\alpha) = K^*y^\delta, \tag{2.12}$$

$$(K^*K + \alpha I)x'(\alpha) = -x(\alpha), \tag{2.13}$$

$$(K^*K + \alpha I)x''(\alpha) = -2x'(\alpha). \tag{2.14}$$

We are now ready to describe the third-order iteration method for solving Eq. (2.10) in details.

Chebyshev method. Given an initial guess α_0, generate the Chebyshev sequence $\{\alpha_n\}_{n=0}^{\infty}$, by

$$\alpha_{n+1} = \alpha_n - \left(1 + \frac{1}{2}L(\alpha_n)\right)\big[G'(\alpha_n)\big]^{-1} G(\alpha_n),$$

where

$$L(\alpha_n) = G(\alpha_n)\big[G'(\alpha_n)\big]^{-2} G''(\alpha_n). \tag{2.15}$$

Halley method. Given an initial guess α_0, generate the Halley sequence $\{\alpha_n\}_{n=0}^{\infty}$, by

$$\alpha_{n+1} = \alpha_n - \left(1 + \frac{1}{2}L(\alpha_n)\left[1 - \frac{1}{2}L(\alpha_n)\right]^{-1}\right)\big[G'(\alpha_n)\big]^{-1} G(\alpha_n),$$

where $L(\alpha_n) = G(\alpha_n)[G'(\alpha_n)]^{-2} G''(\alpha_n)$.

Super-Halley method. Given an initial guess α_0, generate the Super-Halley sequence $\{\alpha_n\}_{n=0}^{\infty}$, by

$$\alpha_{n+1} = \alpha_n - \left(1 + \frac{1}{2}L(\alpha_n)\left[1 - L(\alpha_n)\right]^{-1}\right)\left[G'(\alpha_n)\right]^{-1}G(\alpha_n),$$

where $L(\alpha_n) = G(\alpha_n)[G'(\alpha_n)]^{-2}G''(\alpha_n)$.

The following lemma states a general principle for analyzing the convergence order of a given iterative method.

Lemma 2.5. *Let $\alpha_{n+1} = \phi(\alpha_n)$ represent an iterative process and assume that $\phi^{(p)}(\alpha)$ $(p = 1, 2, \ldots)$ is continuous in a neighborhood of the solution α^*. If*

$$\phi'(\alpha^*) = \phi''(\alpha^*) = \cdots = \phi^{(p-1)}(\alpha^*) = 0, \quad \phi^{(p)} \neq 0,$$

we say that the iterative method is of pth-order convergence.

Now, we estimate the convergence order of the three iterative methods mentioned above.

Theorem 2.6. *Chebyshev method, Halley method and Super-Halley method are all at least of third-order convergence.*

Proof. We first establish the convergence order of the Chebyshev method. The corresponding iteration function reads

$$\phi(\alpha) = \alpha - \left(1 + \frac{1}{2}L(\alpha)\right)\left[G'(\alpha)\right]^{-1}G(\alpha).$$

In terms of the smoothness of the function $G(\alpha)$, the iteration function ϕ is sufficiently smooth.

Assume $G(\alpha^*) = 0$. Observing the definition of $L(\alpha)$ in Eq. (2.15), we have $L(\alpha^*) = 0$ and

$$\frac{d}{d\alpha}L(\alpha) = \left[G'(\alpha)\right]^{-1}G''(\alpha) - 2G(\alpha)\left[G'(\alpha)\right]^{-3}\left[G''(\alpha)\right]^2 + G(\alpha)\left[G'(\alpha)\right]^{-2}G'''(\alpha).$$

Therefore,

$$\begin{aligned}
\phi'(\alpha^*) &= 1 - \left[1 + \frac{1}{2}L(\alpha^*)\right] + G(\alpha^*)\left[G'(\alpha^*)\right]^{-2}G''(\alpha^*)\left[1 + \frac{1}{2}L(\alpha^*)\right] - G(\alpha)\left[G'(\alpha^*)\right]^{-1}\left[\frac{1}{2}\frac{d}{d\alpha}L(\alpha)\right]\Bigg|_{\alpha=\alpha^*} \\
&= \frac{3}{2}\left[L(\alpha^*)\right]^2 + \frac{1}{2}L(\alpha^*) - \frac{1}{2}\left[G'(\alpha^*)\right]^{-3}G'''(\alpha^*)G^2(\alpha^*) - \frac{1}{2}\left[G'(\alpha^*)\right]^{-3}G''(\alpha^*)G'(\alpha^*)G(\alpha^*) \\
&= \frac{3}{2}\left[L(\alpha^*)\right]^2 - \frac{1}{2}\left[G'(\alpha^*)\right]^{-3}G'''(\alpha^*)G^2(\alpha^*) \\
&= 0, \\
\phi''(\alpha^*) &= 3L(\alpha^*)\frac{d}{d\alpha}L(\alpha)\Big|_{\alpha=\alpha^*} + \frac{3}{2}\left[G'(\alpha^*)\right]^{-4}G'''(\alpha^*)G^2(\alpha^*)G''(\alpha^*) \\
&\quad - \frac{1}{2}\left[G'(\alpha^*)\right]^{-3}G^{(4)}(\alpha^*)G^2(\alpha^*) - \left[G'(\alpha^*)\right]^{-2}G'''(\alpha^*)G(\alpha^*) \\
&= 0.
\end{aligned}$$

In a similar way we obtain

$$\phi'''(\alpha^*) = 3\left[\frac{d}{d\alpha}L(\alpha)\Big|_{\alpha=\alpha^*}\right]^2 - \left[G'(\alpha^*)\right]^{-1}G'''(\alpha^*) = 3\left[G'(\alpha^*)\right]^{-2}\left[G''(\alpha^*)\right]^2 - \left[G'(\alpha^*)\right]^{-1}G'''(\alpha^*). \tag{2.16}$$

Thus, the Chebyshev method is at least of third-order convergence.

Next we discuss Halley method. The Super-Halley method can be treated in a similar way. The iteration function of Halley method is

$$\phi(\alpha) = \alpha - \left(1 + \frac{1}{2}L(\alpha)\left[1 - \frac{1}{2}L(\alpha)\right]^{-1}\right)\left[G'(\alpha)\right]^{-1}G(\alpha).$$

Thus,

$$\begin{aligned}
\phi'(\alpha^*) &= 1 - \left[G'(\alpha^*)\right]^{-1}G(\alpha^*)\frac{d}{d\alpha}\left(1 + \frac{1}{2}L(\alpha)\left[1 - \frac{1}{2}L(\alpha)\right]^{-1}\right)\Bigg|_{\alpha=\alpha^*} - 1 - \frac{1}{2}L(\alpha^*)\left[1 - \frac{1}{2}L(\alpha^*)\right]^{-1} \\
&\quad + \left(1 + \frac{1}{2}L(\alpha^*)\left[1 - \frac{1}{2}L(\alpha^*)\right]^{-1}\right)L(\alpha^*) \\
&= 0,
\end{aligned}$$

$$\begin{aligned}
\phi''(\alpha^*) &= -\big[G'(\alpha^*)\big]^{-1} G(\alpha^*)\frac{d^2}{d\alpha^2}\left(1+\frac{1}{2}L(\alpha)\left[1-\frac{1}{2}L(\alpha)\right]^{-1}\right)\bigg|_{\alpha=\alpha^*} - \frac{d}{d\alpha}\left(1+\frac{1}{2}L(\alpha)\left[1-\frac{1}{2}L(\alpha)\right]^{-1}\right)\bigg|_{\alpha=\alpha^*} \\
&\quad + L(\alpha^*)\frac{d}{d\alpha}\left(1+\frac{1}{2}L(\alpha)\left[1-\frac{1}{2}L(\alpha)\right]^{-1}\right)\bigg|_{\alpha=\alpha^*} - \frac{1}{2}L(\alpha^*)\frac{d}{d\alpha}\left[1-\frac{1}{2}L(\alpha)\right]^{-1}\bigg|_{\alpha=\alpha^*} \\
&\quad - \frac{1}{2}\left[1-\frac{1}{2}L(\alpha^*)\right]^{-1}\frac{d}{d\alpha}L(\alpha)\bigg|_{\alpha=\alpha^*} + L(\alpha^*)\frac{d}{d\alpha}\left(1+\frac{1}{2}L(\alpha)\left[1-\frac{1}{2}L(\alpha)\right]^{-1}\right)\bigg|_{\alpha=\alpha^*} \\
&\quad + \left(1+\frac{1}{2}L(\alpha^*)\left[1-\frac{1}{2}L(\alpha^*)\right]^{-1}\right)\frac{d}{d\alpha}L(\alpha)\bigg|_{\alpha=\alpha^*} \\
&= -\frac{d}{d\alpha}\left(1+\frac{1}{2}L(\alpha)\left[1-\frac{1}{2}L(\alpha)\right]^{-1}\right)\bigg|_{\alpha=\alpha^*} \\
&\quad - \frac{1}{2}\left[1-\frac{1}{2}L(\alpha^*)\right]^{-1}\frac{d}{d\alpha}L(\alpha)\bigg|_{\alpha=\alpha^*} + \left(1+\frac{1}{2}L(\alpha^*)\left[1-\frac{1}{2}L(\alpha^*)\right]^{-1}\right)\frac{d}{d\alpha}L(\alpha)\bigg|_{\alpha=\alpha^*} \\
&= 0.
\end{aligned}$$

Similarly we have

$$\phi'''(\alpha^*) = \frac{3}{2}\big[G'(\alpha^*)\big]^{-2}\big[G''(\alpha^*)\big]^2 - \big[G'(\alpha^*)\big]^{-1}G'''(\alpha^*).$$

Thus, the Halley method is at least of third-order convergence. □

Remark 2.7. In general we may have $\phi'''(\alpha^*) \neq 0$. Then the Chebyshev method, the Halley method and the Super-Halley method all have exactly third-order convergence. Otherwise, they may exhibit an even higher order of convergence.

Now we compare the numerical computation costs of the standard Newton method and the cubically convergent methods studied in Theorem 2.6. We will use the total number of multiplications and divisions as a measure for these costs. In order to apply the cubically convergent methods, we need to solve Eqs. (2.12) to (2.14) which are all linear algebraic equations described by the same coefficient matrix. We apply the LU factorization method to solve these linear equations. This means that we only need to calculate the Cholesky factorization once and then carry out three backward eliminations. When applying the cubically convergent methods, the total computation cost in one iteration step is $\frac{n^3}{6} + \frac{5n^2}{2} + \frac{41n}{6}$, which is just $\frac{n^2}{2} + \frac{21n}{6}$ more than the cost for one step of the standard Newton method. As the problem scale becomes large, the computation costs of the three cubically convergent methods and the standard Newton method are nearly the same. On the other hand, the cubically convergent methods converge faster than Newton method, which means that the former require few iteration steps for the approximate solutions to reach the same error tolerance.

3. Numerical experiments

In this section, we apply the cubically convergent iterative methods discussed in previous section to an elliptic equation. Consider the following two-point boundary value problem for the elliptic equation studied in [14]:

$$-\big(q(x)u_x\big)_x = f(x) \quad \text{in } (0,1) \quad \text{with } u(0)=u(1)=0. \tag{3.1}$$

We take the coefficient function $q(x)$ and the observation data z of u as

$$q(x) = e^{1+x^2}, \qquad z = u(f^*) = e^{-x}\sin(\pi x),$$

and the source term $f(x)$ has the form

$$f^* = -q_x e^{-x}\big\{\pi\cos(\pi x) - \sin(\pi x)\big\} + q e^{-x}\big\{2\pi\cos(\pi x) + (\pi^2-1)\sin(\pi x)\big\}.$$

We assume that the available observed data are the superposition of the error free data z and the sinusoidal noise:

$$z^\delta(x) = z(x) + \delta^*\sin\big(1.5\pi(2x-1)\big).$$

In our numerical implementation, we choose the piecewise linear finite element method to solve Eq. (3.1) and the variational equation (2.1). We first partition the domain $\Omega = (0,1)$ into N equally distributed subintervals and then define V^h to be the continuous piecewise linear finite element space with respect to this partition, with $h = 1/N$. Let V_0^h be the subspace of V^h with functions vanishing at the two endpoints $x=0$ and $x=1$. Then the finite element approximation $f_h(\alpha)$ can be obtained by finding $f_h(\alpha) \in V^h$ such that

$$\big(u_h\big(f_h(\alpha)\big), u_h(g)\big) + \alpha\big(f_h(\alpha), g\big) = \big(z^\delta, u_h(g)\big) \quad \text{for all } g \in V^h,$$

Table 1
Optimal α values and the $\alpha's$ obtained by principle.

δ^*	0.002	0.004	0.006	0.008
α_M	0.341×10^{-8}	0.103×10^{-7}	0.178×10^{-7}	0.253×10^{-7}
α_{opt}	0.112×10^{-8}	0.323×10^{-8}	0.778×10^{-8}	0.154×10^{-7}

Table 2
Convergence of the Newton method with $\alpha_0 = 10^{-4}$.

δ^*	0.002	0.004	0.006	0.008
α_n	0.341×10^{-8}	0.103×10^{-7}	0.178×10^{-7}	0.253×10^{-7}
Iter	9	8	7	7

Table 3
Convergence of the Chebyshev method with $\alpha_0 = 10^{-4}$.

δ^*	0.002	0.004	0.006	0.008
α_n	0.341×10^{-8}	0.103×10^{-7}	0.178×10^{-7}	0.253×10^{-7}
Iter	7	6	6	5

Table 4
Convergence of the Halley method with $\alpha_0 = 10^{-4}$.

δ^*	0.002	0.004	0.006	0.008
α_n	0.341×10^{-8}	0.103×10^{-7}	0.178×10^{-7}	0.253×10^{-7}
Iter	6	6	5	5

Table 5
Convergence of the Super-Halley method with $\alpha_0 = 10^{-4}$.

δ^*	0.002	0.004	0.006	0.008
α_n	0.341×10^{-8}	0.103×10^{-7}	0.178×10^{-7}	0.253×10^{-7}
Iter	6	5	5	5

where $u_h = u_h(f_h(\alpha)) \in V_0^h$ satisfies

$$\big(q(x)(u_h)_x, v_x\big) = \big(f_h(\alpha), v\big) \quad \text{for all } v \in V_0^h.$$

In Table 1, the parameter α_{opt} stands for the optimal α-value which achieves the minimum for $\|f(\alpha) - f^*\|_{L^2(\Omega)}$. It can be obtained as follows: we first calculate the L^2-norm error for 100 uniformly distributed α-values in the interval $[10^{-7}, 10^{-5}]$ to find an approximate optimal α, denoted by $\tilde{\alpha}$. Then a much smaller interval including $\tilde{\alpha}$ is chosen to calculate an accurate α_{opt}. The parameter α_M stands for the solution of the general Morozov equation (2.10). We obtain it by using the bisection algorithm. We see that the Morozov discrepancy principle yields a very accurate approximation to the optimal value of α for the considered example.

In Tables 2–5, we present numerical results for different iterative methods using $N = 20$. All the computations are carried out with initial value $\alpha_0 = 10^{-4}$, and the stopping criterion for the iterative methods is chosen as $|\alpha_{i+1} - \alpha_i| \leqslant 10^{-12}$. “Iter” denotes the number of iterations of the specified algorithm achieving the listed α values.

In Table 2, we present the numerical approximation solution of α and the corresponding iteration steps for Newton method with different perturbation scale of δ^*.

Tables 3–5 show the convergence of the Chebyshev method, the Halley method and the Super-Halley method, respectively, with the same initial guess $\alpha_0 = 10^{-4}$, and they show the numbers of iterations of the third-order methods.

References

[1] J. Baumeister, Stable Solution of Inverse Problems, Adv. Lectures Math., Friedr. Vieweg & Sohn, Braunschweig, Germany, 1987.

[2] J. Cheng, M. Yamamoto, One new strategy for a priori choice of regularizing parameters in Tikhonov's regularization, Inverse Problems 16 (4) (2000) 31–38.

[3] J.E. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice Hall Series in Comput. Math., Prentice Hall Inc., Englewood Cliffs, NJ, 1983.

[4] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Math. Appl., vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

[5] H.W. Engl, K. Kunisch, A. Neubauer, Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems, Inverse Problems 5 (4) (1989) 523–540.

[6] H.W. Engl, J. Zou, A new approach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction, Inverse Problems 16 (6) (2000) 1907–1923.

[7] F. González-Gascón, D. Peralta-Salas, Attractors and symmetries of vector fields: The inverse problem, J. Math. Anal. Appl. 335 (2) (2007) 789–807.
[8] C.W. Groetsch, Inverse Problems in the Mathematical Sciences, Vieweg Math. Sci. Eng., Friedr. Vieweg & Sohn, Braunschweig, Germany, 1993.
[9] J.M. Gutiérrez, M.A. Hernández, A family of Chebyshev–Halley type methods in Banach spaces, Bull. Austral. Math. Soc. 55 (1997) 113–130.
[10] P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, SIAM Rev. 34 (4) (1992) 561–580.
[11] K. Ito, K. Kunisch, On the choice of the regularization parameter in nonlinear inverse problems, SIAM J. Optim. 2 (3) (1992) 376–404.
[12] A. Kirsch, An Introduction to the Mathematical Theory of Inverse Problems, Appl. Math. Sci., vol. 120, Springer-Verlag, New York, 1996.
[13] K. Kunisch, On a class of damped Morozov principles, Computing 50 (3) (1993) 185–198.
[14] K. Kunisch, J. Zou, Iterative choices of regularization parameters in linear inverse problems, Inverse Problems 14 (5) (1998) 1247–1264.
[15] A.K. Louis, Inverse problems in medicine, in: Applications of Mathematics in Industry and Technology, Siena, 1988, Teubner, Stuttgart, 1989, pp. 277–287.
[16] V.A. Morozov, Methods for Solving Incorrectly Posed Problems, Springer-Verlag, New York, 1984, translated from Russian by A.B. Aries, translation edited by Z. Nashed.
[17] C.T. Shieh, V.A. Yurko, Inverse nodal and inverse spectral problems for discontinuous boundary value problems, J. Math. Anal. Appl. 347 (1) (2008) 266–272.
[18] A.N. Tikhonov, V.Y. Arsenin, Solutions of Ill-posed Problems, V.H. Winston & Sons, Washington, DC–New York, 1977.
[19] Y.F. Wang, T.Y. Xiao, Fast realization algorithms for determining regularization parameters in linear inverse problems, Inverse Problems 17 (2) (2001) 281–291.
[20] J.L. Xie, J. Zou, An improved model function method for choosing regularization parameters in linear inverse problems, Inverse Problems 18 (3) (2002) 631–643.